DOCUMENT RESUME

ED 167 147 IR 006 905

AUTHOR Lawrence, Barbara
TITLE The Impact of Computers on the Retrieval and Utilization of Chemical Information.
PUB DATE 12 Jun 78
NOTE 38p.; Paper presented at the Annual Meeting of the Special Libraries Association (Kansas City, Missouri, June 12, 1978)

EDRS PRICE MF-$0.83 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Chemical Industry; *Data Bases; *Data Processing; *Information Retrieval; *Information Systems; Visual Aids

ABSTRACT

The use of computers in retrieving bibliographic chemical information is traced through the SDI, batch, and online modes, and related changes are noted in such areas as data base availability, cost, software, and amount of user control. The impact of these changes on both the quality and quantity of chemical information use is discussed, as well as the positive impact on the chemical information professional whose status has been increased by the new skills required, challenges, and technical interactions with users. Cues for the 16 slides which illustrated this presentation are provided in the text, and the visuals for the slides are attached. (Author/RAO)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED167147

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE IMPACT OF COMPUTERS ON THE RETRIEVAL AND UTILIZATION OF CHEMICAL INFORMATION

by

Barbara Lawrence

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
Barbara Lawrence

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

IR006 905

ABSTRACT

THE IMPACT OF COMPUTERS ON THE RETRIEVAL AND UTILIZATION OF CHEMICAL INFORMATION

The computer has had an enormous impact on the retrieval and utilization of chemical information. As I trace the use of computers in retrieving bibliographic chemical information through the SDI, batch, and now online modes, I will note the related changes in such areas as data base availability, cost, software, and amount of user control. These changes have certainly had an impact on both the quality and quantity of chemical information use and indeed have had a positive impact on the chemical information professional. The increased skills required, challenges, and technical interactions with users have increased the status of chemical information workers. The computer has helped us, too.

Barbara Lawrence

EXXON RESEARCH + ENGINEERING CO
PO BOX 121
LINDEN, NJ

IR0006905

Rep Rec 14/34/78

REVISED PROGRAM FOR THE CHEMISTRY DIVISION, SPECIAL LIBRARIES ASSOCIATION ANNUAL MEETING. JUNE 12, 1978, KANSAS CITY, MISSOURI

JOINT SYMPOSIUM SPONSORED BY THE CHEMISTRY DIVISION OF SPECIAL LIBRARIES ASSOCIATION AND SIG BC OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE.

9AM-

ABSTRACTING AND INDEXING SERVICES IN CHEMISTRY AND ALLIED DISCIPLINES. TONI CARBO BEARMAN, NFAIS.

THE PROPOSED CHEMICAL ABSTRACTS SERVICE SUBSTRUCTURE SEARCH SYSTEM. NICK A. FARMER, CHEMICAL ABSTRACTS SERVICE.

THE PRESENT UTILITY OF STRUCTURE-ACTIVITY CORRELATION METHODS - A CREATIVE USE OF INFORMATION. PAUL CRAIG, FRANKLIN INSTITUTE RESEARCH LABORATORIES.

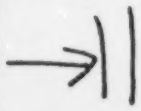
RECENT DEVELOPMENTS IN CHEMICAL NOMENCLATURE. O. C. DERMER, OKLAHOMA STATE UNIVERSITY.

LUNCH TICKETED EVENT
THE FUTURE. RALPH ODETTE, CHEMICAL ABSTRACTS SERVICE.

2PM-

RESPONSE ACTIVITIES AT THE TOXICOLOGY INFORMATION RESPONSE CENTER. E. M. WATERS AND H. B. GERSTNER, TOXICOLOGY INFORMATION RESPONSE CENTER, OAK RIDGE NATIONAL LABORATORY.

ON-LINE NUMERIC DATA BASES. RICHARD C. SIMON, BARRELLE COLUMBUS LABORATORIES.



THE IMPACT OF COMPUTERS ON THE RETRIEVAL AND UTILIZATION OF CHEMICAL INFORMATION. BARBARA LAWRENCE, EXXON RESEARCH AND ENGINEERING COMPANY.

PANEL DISCUSSION:
WALTER GRATTIDGE, CLARKSON COLLEGE
ARLENE SOMERVILLE, UNIVERSITY OF ROCHESTER
DOROTHY LUTZ, XEROX CORPORATION

THE IMPACT OF COMPUTERS ON THE
RETRIEVAL AND UTILIZATION OF CHEMICAL INFORMATION

I. Introduction - Ways The Computer Has Been Used

I've been asked to reflect today on the impact that computers have had on the retrieval and utilization of chemical information. I'll be focusing on the retrieval and utilization of bibliographic chemical information, but won't quite restrict myself to that.

If all I had to do was answer the question, "Have computers had an impact on information retrieval for the chemist?," I could answer that succinctly, "You bet!", and we could finish this talk right now. But there are a few other thoughts I'd like to share with you, and I hope you will find the time spent worthwhile.

First let me itemize the major ways we chemical information specialists (scientists, librarians, researchers) have used computers:

SLIDE 1

1. We have used them for the production of search products. The computer-assisted techniques such as editing, photocomposition, and index sorting and printing have been in extensive use since the mid-60's.
2. We have used them for indexing--automatic indexing, thesaurus production, translation--and for doing research in indexing.

3. Once computers had been used to produce data bases, these data bases were, ipso facto, in computer readable form and became a user tool--we could search them.
4. The first computer searching on these computer-produced data bases was for current awareness, SDI.
5. This created a demand to use the strength of computers for retrospective searching, which was first accomplished in batch mode.
6. Then the computer and telecommunications experts gave us something miraculous to work with--online interactive searching. When we first heard about online searching, it seemed far away--too expensive, a dream. But technology surprised us, and one day (late 1972) it was here, at a price we could afford. Cuadra[1] and others predicted that scientists would be the primary users, searching the literature themselves. So far it hasn't worked out that way. On the contrary, the role of the information intermediary has been enhanced.

   Another thought about online searching: This tool has not replaced all others. For one thing, costs prohibit ready world-wide access to online vendor computers. And the need for computer assistance in other information activities like SDI, resource sharing, and efficient production operation has continued.

SLIDE 2 7. While all this searching went on, the computer was put to work on some non-bibliographic information retrieval problems. The first of these I'll call chemical notation. This covers the broad area relating to chemical structure, such as structure coding, structure activity correlations, and nomenclature standardization for index production at Chemical Abstracts. Fortunately for me these have already been discussed today.

8. There is now a new step that is emerging along the path of progress in chemical information retrieval: the use of computers to handle numeric data. I won't concern myself with raw data capture, or manipulation--but with data accessibility. These data bases take two forms, either as the actual numbers or as data tags in bibliographic files. The visibility of computer information retrieval, and of the retrieval specialist, has pushed the library/information center community toward involvement with these data retrieval tools.

9. Information specialists have used computers in one other significant application--evaluation of information systems and services. Computers have enabled us to design more sophisticated evaluations thereby improving our understanding of information

systems, and user requirements. Since the goal of any evaluation is to produce better products or services for the chemical-information user, I did not want to slight this particular application.

SLIDE 3 Thus, computers have had quite an effect on the retrieval of chemical information, particularly regarding:

accessibility

quantity

qua ity.

I like to think that there has been improvement in these characteristics thanks to the use of computers. Some would argue otherwise.

One effect that I hope we won't argue about is the impact of these computer-related processes on the chemical information specialist. Actually, Hayes[2] has described the 125 year old relationship between libraries and information technology. However, I feel that the current changes in libraries caused by computers are more universal than some previous developments--which were seized upon by those interested in using them, but disregarded by many.

The explosion of online searching into the repetoire of the retrieval specialist enhanced the image of that specialist.[3] This new search tool increased service efficiency, timeliness, and flexibility. References services were expanded to full scale literature-search services. The search process became more visible; customers became consciously

involved in that process. They developed an appreciation of the skills required for effective information retrieval, be they experience, training, subject familiarity, knowledge of Boolean logic, or the ability to recognize serendipity.

However, lest we chemical information specialists become complacent while enjoying our new found status, I'd like to warn that life is getting more complicated. (Even more so than some of us foresaw at a symposium on using data bases in combination held at ASIS mid-year meeting in 1976.[4])

Thanks to the computer, we are being pushed to become information managers, systems analysts,--and better chemists. Online data bases are becoming more complex--and beginning to include data retrieval and manipulation capabilities (data tags or actual numbers).

SLIDE 4 COMPLEX DB	NUMERIC DB
CASIA	PREDICASTS FAMILY
DERWENT/WPI	METALLURGICAL THERMODYNAMIC DATA BASE
PREDICASTS	CHEMTRAN
BIOSIS	MASS SPECTRAL SEARCH SYSTEM

The increasing accessibility of numeric data bases has generated two conflicting thrusts. One is to simplify the retrieval software so that any chemist can use these tools. The other is to bring the systems into the library which already has a communication terminal and online searching experience. Using these data bases will not be easy for many of us. I have no predictions on the outcome of this development, but ASIS SIG/NDB

(Numeric Data Bases), and CODATA (Committee on Data for Science and Technology), among other groups, are actively discussing the issue. And there is a compromise position. We, the chemical information people, can provide the resources, can be the aggressor in bringing useful information to the chemist in a variety of ways. The library is, after all, a resource center. We know how to provide a book catalog and how to teach people to use it, why not do the same with data services?

II. Chronological Trends in Use of Computers in Chemical Information

These are my personal impressions, but what does the literature have to say about computers and chemical information retrieval? Having once been an online searcher, I thought I'd use our computer terminal to search the literature, and see if it supported my opinions on the trends in computer use in chemical information. I searched several data bases for articles on the use of computers; CHEMCON, LISA, SCISEARCH, and retrieved a list of articles from the Journal Chemical Information and Computer Sciences. Then I classified the articles according to the following list:

SLIDE 5

numeric data
searching--online
searching--batch
SDI
notation
production
indexing
overview/management

and calculated the percentage of papers for each subject by year. Some articles were classified under two of these terms. Looking at the results by year, I was dismayed to find no nice straight lines. For one thing the data might be more demonstrative if it went back a few more years (one notable problem of online searching). Further more, the need to improve continuing information dissemination activities doesn't go away, even when that activity is no longer the hot topic.

However, having done all that work, I'll show you what I found.

SLIDE 6 Searching--Online

This does indicate an increase in papers published over the past few years. In 1976 they were a healthy chunk of the whole of publications on computer-assisted information retrieval in chemistry.

SLIDE 7 Searching--Batching

The emphasis on batch searching has indeed diminished. LISA, which covers more foreign literature than the other sources, has more references to batch search systems.

SLIDE 8 Numeric Data Retrieval

There does seem to be increased interest in numeric data retrieval by computer. From what I've seen of this year's literature, I expect that the 1978 figures will continue for use to show growing activity in this field.

SLIDE 9 SDI

Perhaps the peak of publication on SDI services has passed, but it is still an active, and significant function of chemical information centers.

SLIDE 10 Production

The interest in more efficient, more user oriented, and cheaper information products naturally doesn't lag. Publication activity is pretty steady.

SLIDE 11 Notation

The subject of chemical notation has been of interest for a long, long time. I believe computers were used with chemical coding systems before they were used for the retrieval of bibliographic information. Scientists still have much to learn from the relationships between structure and properties, and I think that computers will continue to be important in understanding these relationships.

III. Impact of Computers On Bibliographic Retrieval Techniques

I'd like to bring the focus of this back to the way computers have been used to retrieve bibliographic chemical information. I'll discuss the developments as seen by the user roughly in chronological order.

- Production of Data Bases (Secondary Services) SLIDE 12

  Computers used to: Print (photocompose)
  Organize
  Edit
  Produce Subsets
  Share resources with other DB producers.

  Result: Improve timeliness of service
  Hold down costs
  Provide flexible, targeted services

- SDI Services (Personalized current awareness) SLIDE 13

  DB availability: In-house for a small number of DB's
  Information Center - larger selection

  Cost/Profile: Reasonable, $100-$300/year

  Software: Had more flexibility than retrospective search programs since smaller amounts of data were processed. Right and

	left truncation (important for chemical nomenclature). Serial search process.
User Control	yes, in profile definition--but depend on skills of intermediary. Iterative process. Sometimes limited by software, data base availability.

- Batch Retrospective Search — SLIDE 14

DB Availability	Very few firms could offer this in-house. Use of outside centers--more data bases.
Cost/Search	Expensive ($150-$300/search), for the search--not including information specialists cost. Only use for major projects.
Software	Serial at first, later disc. At first had limited capability (some limits in DB design).
User Control	Decreased because of remote processing--more intermediaries. Cost limited experience.

- Online Interactive — SLIDE 15

DB Availability	Explosion, 10 in 1973, 301 in 1975, 362 in 1977. DB Characteristics - all fields searchable; complex files.
Costs	Sharing reduces costs (can search often, big or small). Cooper - $ 28.41[5] Hawkins - 22.00[6] Collette - 112.00[7]
Software	More and more flexibility. All data fields accessible. Easy command language and better user education. But intermediaries must work hard to keep up to date.
User Control	Work as a team. Interactive. Full time availability. Appropriate for various types of queries. Search at start of a project. Locate information on candidate/consultant. Try a "wild idea." Not just for a major project.

IV. Computer Impact on Utilization

Computers have helped make more information more accessible. Increased utilization of information can be demonstrated with online search statistics; Martha Williams[8] has reported that 1 million online searches (1 search = 1 query/data base) were run in 1975, and 2 million in 1977.

SLIDE 16 EXXON RESEARCH AND ENGINEERING COMPANY

	1974	1975	1976	1977
No. Searches	2102	3321	4344	4825
		BELL LABS		
		604	2852	6172

This increase in query volume, is not just a function of increasing numbers of users. It is also a function of the changes in the way information (especially computer-retrieved information) is being used by chemists. Literature searching is no longer an activity reserved for major project startup. Literature searches--big and small--occur at all stages of the R&D process. At Exxon Research, we also use computer searching techniques to identify consultants, analyze technology trends, and provide new employees with review articles in their new fields. Even the regular subject request has become more complex, in tandem with our increased experience and the availability of more complex data bases. Our users no longer ask for information on chemical additives used as anti-wear agents in engine lubes. New they ask about particular chemical classes, or

the activities of a specific company in relation to anti-wear additives. Computers have helped the information specialist and the user to define and locate the information <u>really</u> needed.

V. Conclusion

I have tried to show the ways that computers have indeed had an effect on chemical information retrieval techniques. Improvements in software, hardware, and telecommunications have placed the increasingly computer-produced data bases in systems that are flexible, cost-effective, and thereby accessible. All of this has increased our utilization of chemical information.

In the past few years especially, computer technology has had as much of an impact on the information specialist, as on the generation and retrieval of chemical information, or on the chemical information users. Our image has improved with the improved (and often complex) services we can provide.

One notable characteristic of computers is their speed. Now our customers want us to perform at computer speed--not just in retrieving, references, but in providing the documents, and, in general, responding to all requests instantly.

All of this momentum has put us under pressure. No longer can you be just a reference librarian. As I said earlier, we are being pushed into becoming information managers, marketing experts, systems analysts, and better chemists.

I see all of this as a challenge, not a threat. Chemists do want to get closer to the retrieval process--especially when data rather than citations are involved. And although the information specialist can retrieve, edit, organize, and analyze for the client chemist, not all clients want to be spoon fed.

There are several projects working toward putting the chemist at an information computer terminal--at least for some of his or her retrieval work. Martha Williams has been developing a data base selector; MIT has developed a front end minicomputer command language translator; and Charles Meadows plans to build tutorial programming into the online search process. And there is more work of this nature going on.

Still, I don't think we, the information chemists, will be pushed aside. Herman Skolnick[9] has described two roles for the chemical information scientist in the computer age: The first is to help the chemist build and use computer systems that will retrieve from, browse through, and interact with the literature in a companionable way. And secondly, he says that "computers and telecommunications... are in no way a substitute for the chemical information scientists' exploration of the fundamental logical nature of chemical information."

An 1973 NAS/ARC report, "The Status of Chemical Information"[10] analyzes the strengths of and dissatisfaction with computer systems. This committee felt as do I that the information chemist's role in providing resources, knowing which to use and how, educating users, and in knowing the most effective and economical ways to meet user needs will continue.

"What matters is information utilization, not which hands push the keyboard on a computer terminal. We are the communication link between resource providers, system designers, and users.

But computers act in nanoseconds. So we must work hard to keep up.

REFERENCES

1. Cuadra, Carlos A., On-Line Systems: Promise and Pitfalls, Journal of the American Society for Information Science. 1971; 22(2):107-114.

2. Hayes, Robert M., The Changing Role of the Librarian, Journal of Chemical Documentation. 1974; 14(3):118-120.

3a. Lawrence, B., Weil, B. H., and Graham, M. H., Making On-Line Search Services Available in an Industrial Research Environment, Journal of the American Society for Information Science. 1974; 25:364-369.

b. Wanger, Judith, Cuadra, Carlos A., and Fishburn, Mary, Impact of On-Line Retrieval Services: A Survey of Users, 1974-1975. System Development Corporation, Santa Monica, California.

c. Krentz, David M., On-Line Searching--Specialist Required, Journal of Chemical Information and Computer Sciences. 1978; 18(1):4-9.

4. Use of Data Bases in Combination, SIG/BC Panel Discussion, American Society for Information Science, 5th ASIS Mid-Year Meeting, May 20-22, 1976, Nashville, Tennessee.

5. Cooper, Michael D. and DeWath, Nancy A., The Cost of On-Line Bibliographic Searching, Journal of Library Automation. 1976; 9(3):195-209.

6. Hawkins, Donald T., Impact of On-Line Systems on a Literature Searching Service, Special Libraries. 1976; 67(12):559-567.

7. Collette, A. D. and Price, J. A., A Cost/Benefit Evaluation of Online Interactive Bibliographic Searching in a Research and Engineering Organization, Report.

8. Williams, M., 1977 Data-Base and On-Line Statistics, Bulletin of the American Society for Information Science. 1977; 4(2):21.

9. Skolnick, Herman, Milestones in Chemical Information Science, Journal of Chemical Information and Computer Sciences. 1976; 16(4):187-193.

10. Holm, Bart, Howell, Mary Gertrude, Kennedy, H. Edward, Kuney, Joseph H., and Rush, James E., The Status of Chemical Information, Journal of Chemical Documentation. 1973; 13(4):171-183.

COMPUTER USE IN CHEMICAL INFORMATION

BIBLIOGRAPHIC

PRODUCTION OF SEARCH TOOLS

INDEXING

SDI

BATCH SEARCHING

ONLINE SEARCHING

COMPUTER USE IN CHEMICAL INFORMATION

NON-BIBLIOGRAPHIC

- CHEMICAL NOTATION
- NUMERIC DATA RETRIEVAL
- EVALUATION

SLIDE 2

CHARACTERISTICS OF COMPUTER INFORMATION RETRIEVAL

ACCESSIBILITY OF DATA

QUANTITY OF INFORMATION AVAILABLE

QUALITY OF RETRIEVAL RESULTS

COMPLEX DB	NUMERIC DB
CASIA	PREDICASTS FAMILY
DERWENT/WPI	METALLURGICAL THERMO-DYNAMIC DATA BASE
PREDICASTS	CHEMTRAN
BIOSIS	MASS SPECTRAL SEARCH SYSTEM

NUMERIC DATA

SEARCHING--ONLINE

SEARCHING--BATCH

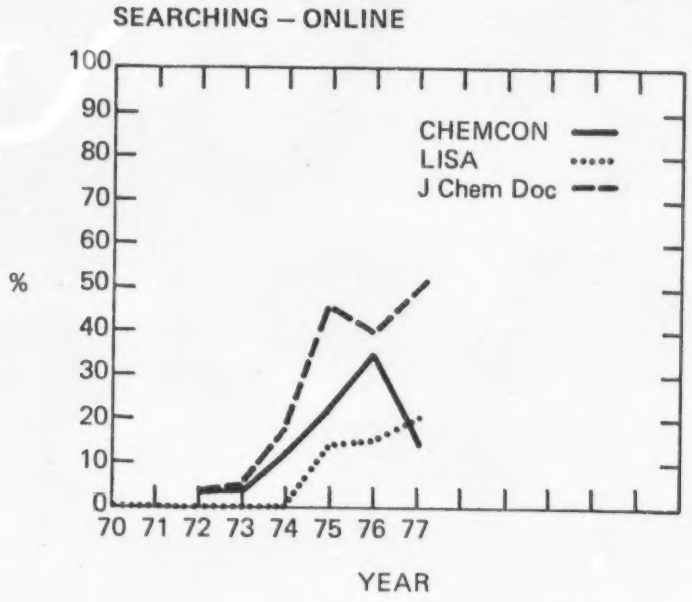
SDI

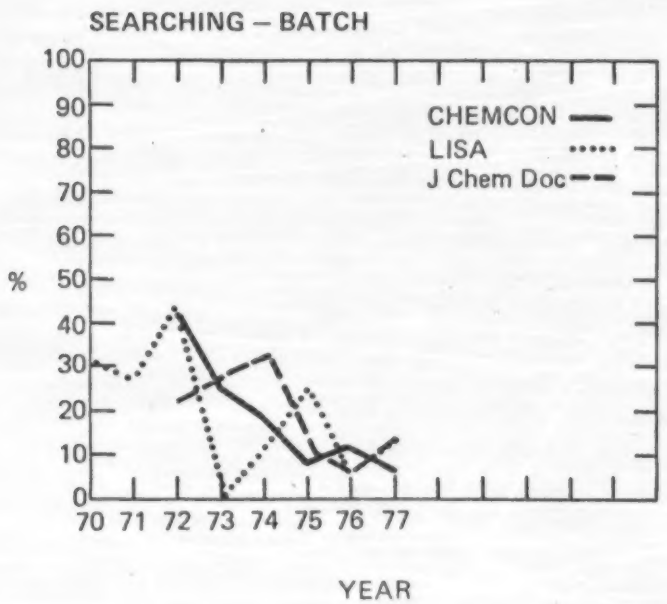
NOTATION

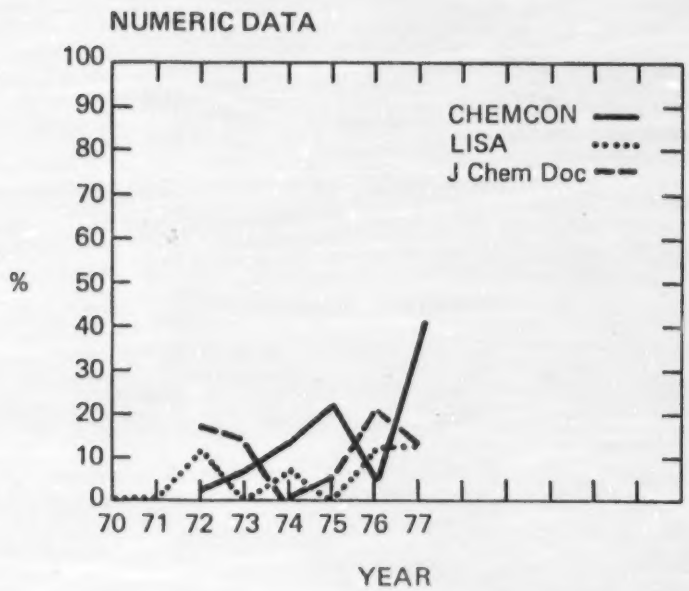
PRODUCTION

INDEXING

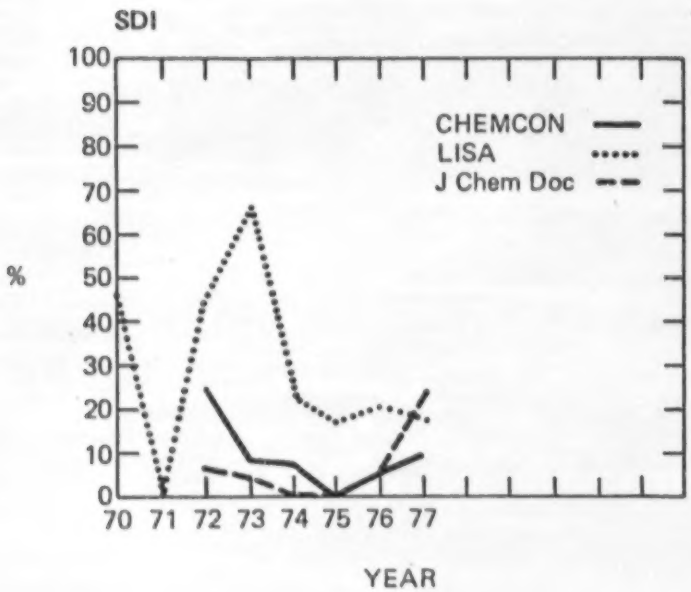
OVERVIEW/MANAGEMENT

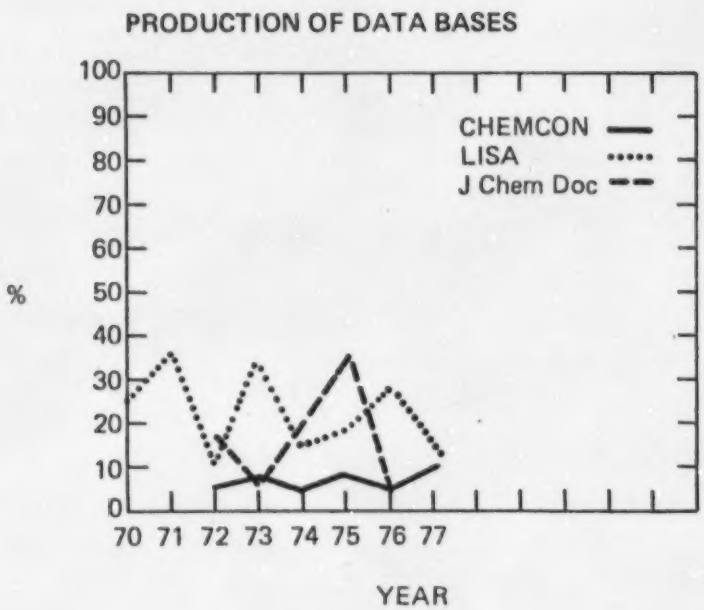

SEARCHING – ONLINE
100
90
80
70
60
%
50
40
30
20
10
0
70 71 72 73 74 75 76 77
YEAR
CHEMCON
LISA
J Chem Doc


SEARCHING – BATCH
100
90
80
70
60
50
40
30
20
10
0
%
CHEMCON
LISA
J Chem Doc
70 71 72 73 74 75 76 77
YEAR

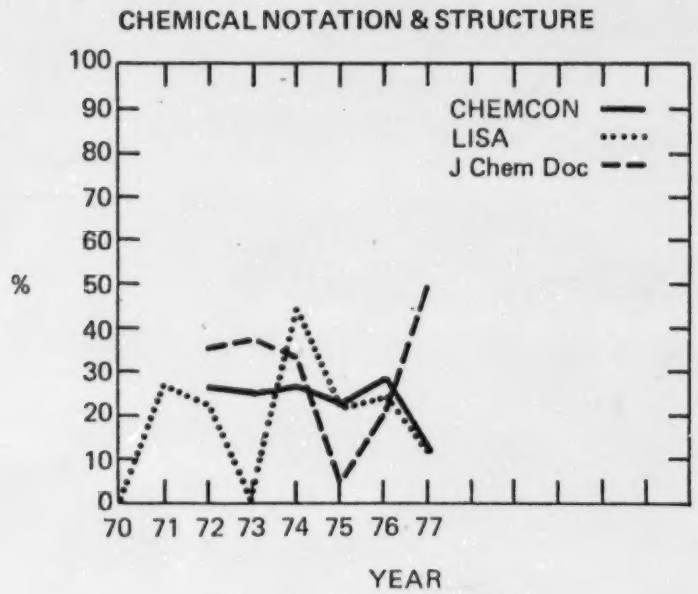

NUMERIC DATA
100
90
80
70
60
50
40
30
20
10
0
%
CHEMCON
LISA
J Chem Doc
70 71 72 73 74 75 76 77
YEAR

8


SDI
100
90
80
70
60
50
40
30
20
10
0
%
70 71 72 73 74 75 76 77
YEAR
CHEMCON
LISA
J Chem Doc


PRODUCTION OF DATA BASES
100
90
80
70
60
50
40
30
20
10
0
%
CHEMCON
LISA
J Chem Doc
70 71 72 73 74 75 76 77
YEAR

10


CHEMICAL NOTATION & STRUCTURE
100
90
80
70
60
50
40
30
20
10
0
%
CHEMCON
LISA
J Chem Doc
70 71 72 73 74 75 76 77
YEAR

DATA BASE PRODUCTION

COMPUTERS - PHOTOCOMPOSE
ORGANIZE
EDIT
COOPERATE

RESULTS - IMPROVE TIMELINESS
MINIMIZE COST
PROVIDE TARGETED SERVICES

SDI

DB AVAILABILITY	:	IN-HOUSE AND SERVICE CENTER
COST	:	REASONABLE
SOFTWARE	:	FLEXIBLE
USER CONTROL	:	PROFILE DEFINITION. LIMITED BY SOFTWARE, SKILL OF INTERMEDIARY, LOCATION OF SERVICES.

BATCH RETROSPECTIVE SEARCH

DB AVAILABILITY:		MOSTLY EXTERNAL SERVICE CENTER
		MORE DATA BASES
COST	:	EXPENSIVE
SOFTWARE	:	IMPROVE FROM SERIAL TO DIRECT ACCESS SYSTEMS
		DATA BASE DESIGN LIMITED AT FIRST
USER CONTROL	:	DECREASED BY REMOTE PROCESSING
		COST LIMITED EXPERIENCE

ONLINE INTERACTIVE SEARCHING

DB AVAILABILITY:		EXPLOSION
COST	:	REDUCED DRASTICALLY
SOFTWARE	:	INCREASED FLEXIBILITY IMPROVED DATA BASE DESIGN BETTER USER EDUCATION
USER CONTROL	:	INCREASED BY INTERACTIVENESS, FULL-TIME AVAILABILITY OF SYSTEMS, APPROPRIATE FOR VARIOUS TYPES OF QUERIES.

EXXON RESEARCH AND ENGINEERING COMPANY

	1974	1975	1976	1977
NO. SEARCHES	2102	3321	4344	4825

BELL LABS

604	2852	6172	